

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: February 23, 2003, 21:05:35 ; Search time 3922 Seconds
(without alignments)
10907.987 Million cell updates/sec

Title: US-09-725-010-1
Perfect score: 1470
Sequence: 1 ggaatcttctttttcttttc.....ctttttgttctgcaaatatg 1470

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 2054640 seqs, 14551402878 residues 4109280
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htg_mus.*
34: em.htg_pln.*
35: em.htg_rod.*
36: em.htg_mam.*
37: em.htg_vrt.*
38: em.sy.*
39: em.htgo_hum.*
40: em.htgo_mus.*
41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1470	100.0	1470	6	AX154641	AX154641 Sequence
2	505	34.4	2682	8	CAU64206	U64206 Candida alb
3	153	10.4	2188	8	AF001978	AF001978 Candida a
c 4	136.2	9.3	42565	8	CAC335A5	AL033396 C.albican
c 5	91.2	6.2	86827	3	PFMAL3P5	AL034556 Plasmodiu
c 6	86.8	5.9	104992	2	AC005504	AC005504 Plasmodiu
c 7	86.8	5.9	169546	2	AC004157	AC004157 Plasmodiu
c 8	75.2	5.1	104014	2	AC116921	AC116921 Dictyoste
c 9	73.4	5.0	14867	3	AE001398	AE001398 Plasmodiu
c 10	73.4	5.0	164399	3	PFMAL3P6	Z98551 Plasmodium
c 11	72.8	5.0	113880	3	PFMAL3P4	AL008970 Plasmodiu
c 12	70.8	4.8	2009	6	AX457067	AX457067 Sequence
c 13	70.4	4.8	192929	2	AC005505	AC005505 Plasmodiu
c 14	69.4	4.7	56099	2	AC115598	AC115598 Dictyoste
c 15	69.2	4.7	253305	3	PFMAL3P7	AL034559 Plasmodiu
c 16	68.2	4.6	115489	2	AC117072	AC117072 Dictyoste
c 17	68	4.6	9810	6	AX345328	AX345328 Sequence
c 18	67.2	4.6	7347	1	AF211124	AF211124 Carsonell
c 19	67.2	4.6	67970	3	PFMAL1P3	AL031746 Plasmodiu
c 20	67	4.6	130349	9	AC011593	AC011593 Homo sapi
c 21	67	4.6	198431	9	AC109994	AC109994 Homo sapi
c 22	66.8	4.5	97371	8	ATT28A8	AL162691 Arabidops
c 23	66.6	4.5	1141	6	AX083744	AX083744 Sequence
c 24	66.4	4.5	14635	3	AE001423	AE001423 Plasmodiu
c 25	66.4	4.5	18624	6	AX346604	AX346604 Sequence
c 26	66.4	4.5	30726	2	AC117269	AC117269 Dictyoste
c 27	66.2	4.5	1192	9	HSA323759	AJ323759 Homo sapi
c 28	66.2	4.5	115489	2	AE001398	AE001398 Plasmodiu
c 29	66.2	4.5	115489	2	AC117072	AC117072 Dictyoste
c 30	66.2	4.5	133148	9	AC009435	AC009435 Homo sapi
c 31	66	4.5	17538	6	AX346059	AX346059 Sequence
c 32	66	4.5	201299	9	AC007158	AC007158 Homo sapi
c 33	65.8	4.5	156060	2	AC004153	AC004153 Plasmodiu
c 34	65.6	4.5	7584	6	AX286935	AX286935 Sequence
c 35	65.4	4.4	22977	9	AC106883	AC106883 Homo sapi
c 36	65.4	4.4	162515	9	AC107032	AC107032 Homo sapi
c 37	65.2	4.4	23676	2	AC115599	AC115599 Dictyoste
c 38	65.2	4.4	124820	2	AC117073	AC117073 Dictyoste
c 39	65.2	4.4	141017	2	AC116962	AC116962 Dictyoste
c 40	65.2	4.4	165260	9	AC024341	AC024341 Homo sapi
c 41	65	4.4	43993	2	AC116965	AC116965 Dictyoste
c 42	65	4.4	165059	9	AC106856	AC106856 Homo sapi
c 43	65	4.4	321003	2	PFMAL4P3	AL035476 Plasmodiu
c 44	64.8	4.4	1141	6	AX083744	AX083744 Sequence
c 45	64.8	4.4	5884	6	AX347067	AX347067 Sequence

ALIGNMENTS

RESULT 1
AX154641
LOCUS AX154641 1470 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 1 from Patent WO0138550.
ACCESSION AX154641
VERSION AX154641.1 GI:14536200
KEYWORDS
SOURCE Candida albicans.
ORGANISM Candida albicans
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE
AUTHORS
TITLE 1 (bases 1 to 1470)
METHODS for altering the expression of hyphal-specific genes
JOURNAL Patent: WO 0138550-A 1 31-MAY-2001;

FEATURES
 source Sundstrom, Paula (US)
 1. 1470
 /organism="Candida albicans"
 /db_xref="taxon:5476"
 BASE COUNT 509 a 260 c 186 g 515 t
 ORIGIN

Query Match 100.0%; Score 1470; DB 6; Length 1470;

Best Local Similarity 100.0%; Pred. No. 26-225;

Matches 1470; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCTTTCTTTTCATTTCCCTTAAACCGCATCAAGAAAGAGTGGAAATAAAGCTAT 60
 Db 1 GGATCTTTCTTTTCATTTCCCTTAAACCGCATCAAGAAAGAGTGGAAATAAAGCTAT 60
 QY 61 GATAAATGTCATTTTGTGTAATCAATCAACTAAGCAGCTTTGACAGTTTAAAGAGTACG 120
 Db 61 GATAAATGTCATTTTGTGTAATCAATCAACTAAGCAGCTTTGACAGTTTAAAGAGTACG 120
 QY 121 TTGTTGTGTCCTCGTCTCGTCTAATTTCTGTGACGAGGATTAAACAGAAATACAG 180
 Db 121 TTGTTGTGTCCTCGTCTCGTCTAATTTCTGTGACGAGGATTAAACAGAAATACAG 180
 QY 181 GAACCCCTCCAAAAAATAATTTTGGACCTTACAGCAGACATAATTCGGGATAAACTTGC 240
 Db 181 GAACCCCTCCAAAAAATAATTTTGGACCTTACAGCAGACATAATTCGGGATAAACTTGC 240
 QY 241 CATATAAAAACTCTTGAACATACGATATGTTATTTCTTTCATCACTGGAATATTTT 300
 Db 241 CATATAAAAACTCTTGAACATACGATATGTTATTTCTTTCATCACTGGAATATTTT 300
 QY 301 GCTTTTTTTTAACTATTGAACAAATGAAAAAAGAAATGAAAAAGTAAAGAGTTGCC 360
 Db 301 GCTTTTTTTTAACTATTGAACAAATGAAAAAAGAAATGAAAAAGTAAAGAGTTGCC 360
 QY 361 TAACCATGAAATTAAGGTAGCTAGCTTTTCCGTGATGCGTTTAACTAAAGAAATAA 420
 Db 361 TAACCATGAAATTAAGGTAGCTAGCTTTTCCGTGATGCGTTTAACTAAAGAAATAA 420
 QY 421 CAAAGTTATTAGCGATAACCTGCGTAAAGTGTCAACAAAATATATTTTGCAGCTTAGCT 480
 Db 421 CAAAGTTATTAGCGATAACCTGCGTAAAGTGTCAACAAAATATATTTTGCAGCTTAGCT 480
 QY 481 CTATAGAAATATACAACTAAATCCCTTAAGAAATTTCTCTATATATAATAGAAATCC 540
 Db 481 CTATAGAAATATACAACTAAATCCCTTAAGAAATTTCTCTATATATAATAGAAATCC 540
 QY 541 CTCTCACAGTGAACCTGAATATCCATCTGAATATATCATCTCACTCAATTCATCAATAA 600
 Db 541 CTCTCACAGTGAACCTGAATATCCATCTGAATATATCATCTCACTCAATTCATCAATAA 600
 QY 601 TAGATTAGTGTATTGTTCTCTCAGTACAATTAACCATATGCAATGCTAGCTATTG 660
 Db 601 TAGATTAGTGTATTGTTCTCTCAGTACAATTAACCATATGCAATGCTAGCTATTG 660
 QY 661 TTCATAATTAGCCATGTTGACACCCCTAATTCGAACATTAACGTATGCCATATTTTCTT 720
 Db 661 TTCATAATTAGCCATGTTGACACCCCTAATTCGAACATTAACGTATGCCATATTTTCTT 720
 QY 721 CTCCTTCTCTGTTTCTTAACAAAATGTTCCAGAAATTTTAAAAAATATTGAAAA 780
 Db 721 CTCCTTCTCTGTTTCTTAACAAAATGTTCCAGAAATTTTAAAAAATATTGAAAA 780
 QY 781 AACACATAACACTTTGAGTATGATAATCAACTATTGACTTGTGTTTGAAGTAAAGAAT 840
 Db 781 AACACATAACACTTTGAGTATGATAATCAACTATTGACTTGTGTTTGAAGTAAAGAAT 840
 QY 841 CAAATTTTTTCTAAGTGAATGACTTTTACATCACTGGAATGATTTTGCATCTAC 900
 Db 841 CAAATTTTTTCTAAGTGAATGACTTTTACATCACTGGAATGATTTTGCATCTAC 900
 QY 901 TACTATAAGCTCAAAACAAATATCTTTCAAAAATGTTTATAATTAAACAGTCACTATAAT 960

Db 901 TACTATAAGCTCAAAACAAATATCTTTCAAAAATGTTTATAATTAAACAGTCACTATAAT 960
 QY 961 TCTTTGGATCCAAAACAAAGAAATTCGGAAATCTGAGGATAAATGTCGACTCACAATTC 1020
 Db 961 TCTTTGGATCCAAAACAAAGAAATTCGGAAATCTGAGGATAAATGTCGACTCACAATTC 1020
 QY 1021 AATTGAAAAAGGAGAGTGTGTTAGGCTCATAAATCGCTTATAATGTAACCTCTAAAGTAA 1080
 Db 1021 AATTGAAAAAGGAGAGTGTGTTAGGCTCATAAATCGCTTATAATGTAACCTCTAAAGTAA 1080
 QY 1081 TCTAAAACAAACACACACCTTTCTAAAACCTATAAATAAACCTAATGGCTCACAACCGG 1140
 Db 1081 TCTAAAACAAACACACACCTTTCTAAAACCTATAAATAAACCTAATGGCTCACAACCGG 1140
 QY 1141 GATAAGTTAGTTAGCCAGCTGTTTTTTTTTTTGGCTTATTTTATGACTACATTTTGT 1200
 Db 1141 GATAAGTTAGTTAGCCAGCTGTTTTTTTTTTTGGCTTATTTTATGACTACATTTTGT 1200
 QY 1201 CACTTTTCTCGACTTTTAATACCGTTTTTTCGAACTTCTCTTTGTATCAGCTGATCCG 1260
 Db 1201 CACTTTTCTCGACTTTTAATACCGTTTTTTCGAACTTCTCTTTGTATCAGCTGATCCG 1260
 QY 1261 CCTTTTTTAACATAGCAACTCTTTGAAAGTCCCTTTTCCCACTATTTTATCATTTCT 1320
 Db 1261 CCTTTTTTAACATAGCAACTCTTTGAAAGTCCCTTTTCCCACTATTTTATCATTTCT 1320
 QY 1321 TGAATATATCTAATCAGATAGTTTTTCAAAAACATAAATAACGGTCAAAATAACCGGCT 1380
 Db 1321 TGAATATATCTAATCAGATAGTTTTTCAAAAACATAAATAACGGTCAAAATAACCGGCT 1380
 QY 1381 ATTTCAATTTCCATTCACACTTTGTTTCTCAACAATATCAACACACAGGAATCTCCTA 1440
 Db 1381 ATTTCAATTTCCATTCACACTTTGTTTCTCAACAATATCAACACACAGGAATCTCCTA 1440
 QY 1441 TAGTCACTCGCTTTTAGTTTCGTCATATG 1470
 Db 1441 TAGTCACTCGCTTTTAGTTTCGTCATATG 1470

RESULT 2

CAU64206

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

CAU64206 2682 bp DNA linear PLN 17-DEC-1998
 Candida albicans hyphal wall protein 1 (HWPI) gene, complete cds.

U64206

U64206.1 GI:4028879

Candida albicans.

Candida albicans.

Saccharomycetales; Ascomycota; Eukaryota; Fungi; Ascomycota; Saccharomycetales; Saccharomycetes;

1 (bases 1 to 2682)

Staab,J.F., Ferrer,C.A. and Sundstrom,P.

Developmental expression of a tandemly repeated, proline-and

glutamine-rich amino acid motif on hyphal surfaces on Candida

albicans

J. Biol. Chem. 271 (11), 6298-6305 (1996)

96198091

8626424

2 (bases 1 to 2682)

Staab,J.F. and Sundstrom,P.

Genetic organization and sequence analysis of the hypha-specific

cell wall protein gene HWPI of Candida albicans

Yeast 14 (7), 681-686 (1998)

98301186

9639315

3 (bases 1 to 2682)

Staab,J.F. and Sundstrom,P.

Direct Submission

Submitted (18-JUL-1996) Medical Microbiology & Immunology, Ohio

State University, 333 West 10th Avenue, Columbus, OH 43210, USA

4 (bases 1 to 2682)

Staab,J.F.

QY	1446	ACTCGCTTTTAGTTTCGCAATAG	1470
Dd	481	ACTCGCTTTTAGTTTCGCAATAG	505
RESULT 3			
AF001978		2188 bp	DNA linear
LOCUS		Candida albicans ECE2 gene, complete cds.	PLN 24-JUL-1997
DEFINITION		Candida albicans ECE2 gene, complete cds.	
ACCESSION		AF001978	
VERSION		AF001978.1	GI:2275335
KEYWORDS		Candida albicans.	
SOURCE		Candida albicans	
ORGANISM		Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; mitosporic Saccharomycetales; Candida.	
REFERENCE		Sharkey J.L., Saporito-Irwin, S.M. and Fonzi, W.A.	
AUTHORS		ECE2, a differentially expressed gene of Candida albicans	
TITLE		Unpublished	
REFERENCE		2 (bases 1 to 2188)	
AUTHORS		Sharkey, J.L., Saporito-Irwin, S.M. and Fonzi, W.A.	
TITLE		Direct Submission	
JOURNAL		Submitted (01-MAY-1997) Microbiology and Immunology, Georgetown University, 3900 Reservoir Road NW, Washington, DC 20007-2197, USA	
COMMENT		On Jul 24, 1997 this sequence version replaced gi:2209195.	
FEATURES		Location/Qualifiers	
source		1..2188	/organism="Candida albicans"
			/strain="SC5314"
			/db_xref="taxon:5476"
			/chromosome="4"
			151..2055
			/gene="ECE2"
			151..2055
			/name="ECE2"
gene			
CDS			

[illegible]

Qy	1184	ATGACTACATTTGGTTTCACCTTTTGGTTGGACCTTTAAATACCGTTTTCGCAACTTCCTCTT	1243			
Db	73897	ATTAATTAATATATATGTTTATATTTTAAATTTAAATAAATAATATATTTTATAAAATTAATTA	73956			
Qy	1244	TGATACACCTGTATCCGCCCTTTTAAACATAGCAACTCTTGTAAAGTCCCTTCTTTCTTCC	1303			
Db	73957	TTTATTAATAAATAAATAAATAAATAAATAATATTTATATAATAATACATAAAATATT	74016			
Qy	1304	CACATATTTATCATCTTGTAAATATGTAATCAGAATAGTTTTTCCAAAACATATAAATAAC	1363			
Db	74017	TATTACTATATATTAATAAATAAAGTATAAAATATCTTCTTAAATTATATTTAA	74076			
Qy	1364	GGTCAAAATAACCGGCTATTTTCAATTTCCATTCGAAGTCTGTTTCTCAACAATATCAAAC	1423			
Db	74077	ATTAATTTAATATATATATATATAATAAATAAATAATATTTATTTAATTAATAATA	74136			
Qy	1424	ACAACAGGAATCTCCTATAGTACGCTGTTTAGTT	1459			
Db	74137	ATATTATATAATATTATATATTGTTTATAATT	74172			
RESULT 7						
LOCUS	AC004157	169546 bp	DNA linear HTG 12-AUG-2000			
DEFINITION	Plasmidium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN PROGRESS ***, 2 unordered pieces.					
ACCESSION	AC004157					
VERSION	AC004157.8	GI:9797712				
KEYWORDS	HTG: HTGS PHASE1.					
SOURCE	Plasmidium falciparum.					
ORGANISM	Plasmidium falciparum.					
REFERENCE	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.					
AUTHORS	1 (bases 1 to 169546) Hyman,R.W., Fung,E.L., Qin,F., Rowley,D., Mao,J., Tanaki,T., Kurdi,O.B., Conway,A.B. and Davis,R.W.					
TITLE	Plasmidium falciparum 3D7 chromosome 12					
JOURNAL	Unpublished					
REFERENCE	2 (bases 1 to 169546) Hyman,R.W., Qin,F., Fung,E.L., Conway,A.B. and Davis,R.W.					
AUTHORS	Direct Submission					
TITLE	Submitted (19-FEB-1998) Stanford DNA Sequencing and Technology					
JOURNAL	Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA					
COMMENT	On Aug 12, 2000 this sequence version replaced gi:8810447. * NOTE: This is a 'working draft' sequence. It currently * consists of 2 contigs. The true order of the pieces * is not known and their order in this sequence record is * arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence * as soon as it is available and the accession number will * be preserved. * 1 23466: contig of 23466 bp in length * 23467 23666: gap of unknown length * 23667 169546: contig of 145880 bp in length.					
FEATURES	Location/Qualifiers					
source	1..169546 /organism="Plasmodium falciparum" /db_xref="taxon:5833" /chromosome="12" /clone="PFYAC293" /clone="3D7"					
BASE COUNT	69871 a 15381 c 15705 g 68389 t 200 others					
ORIGIN						
Query Match 5.9%; Score 86.8; DB 2; Length 169546;						
Best Local Similarity 43.2%; Pred. No. 3.7e-05;						
Matches 612; Conservative 0; Mismatches 797; Indels 7; Gaps 4						
Qy	50	AATAAAGCTATGATAATGTTGATTTGGTGAATTCATCACTAAGCAGCTTTTGACAGT	109			
Db	84342	ATTATAATAATAAATAATATATAATTATTAATAATAATAATATATATAATAATAATAATA	84401			

QY 110 TAAAGAGTACGTTGTTGTCCTCGTCTGCTAAATTTCTGTTGACGAGGATTAATAAC 169
Db 84402 TAATATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 84461
QY 170 AAGAAATACAGCAACCTCCCAAAAAAATTTTGACCTTACACCCACATTAATTCG 229
Db 84462 AATATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 84521
QY 230 GATAAAGTGCATTAATAATAAAGCTTTTGAACATACGATATGTTATCTTTTCATAACT 289
Db 84522 ATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 84581
QY 290 GGAATATTTTGGCTTTTACATTAATGAACAATTTGAAAAAAGGAAATGAAAGG 349
Db 84582 TAATTAATAACAATAATTTAAATAATAATAATAATAATAATAATAATAATAATA 84641
QY 350 TAAGAGTTGCTCAACCAATGAAATAATAAGCTTAAGGTTTTTCCATGCTTAACATA 409
Db 84642 TATAAATTAAGAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 84701
QY 410 AAGGAAATACAAAAGTTATTAGCGATAACCTGCGTAAGGTGTCACAAAAATATATTT 469
Db 84702 AATAAATTAACCAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 84761
QY 470 GCAGTTAGCTCTATAGAAAATATACAACTAAATCCTTAAGGAATTCCTCTATATA 529
Db 84762 ATA-AATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 84820
QY 530 ATAGGAATCCCTCTCACAGTGAATATATCCATCTGAAATATACGTCCCAATATTC 589
Db 84821 TAAATAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 84880
QY 590 CATCAATAAATAAGTAGTATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 649
Db 84881 ATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 84940
QY 650 CTAGCTTATTTGTCATAATTTAGCCATGTCACACACCTTAATTCGAAATTAATCTG 709
Db 84941 AATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 85000
QY 710 ATATTTTCTGTCCT 767
Db 85001 ATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 85060
QY 768 AATATTTGAAAAACACATAACACTTTGAGTATGATATTAATCAACTATGACTTTG 827
Db 85061 ATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 85120
QY 828 GAAAGTAAAGAAATCAAAATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 887
Db 85121 TAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 85180
QY 888 TATTGCACTACTACT--ATAAGCTCAACAAATATATCTTTCAAAAATGTTATATAATA 945
Db 85181 TAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 85240
QY 946 CAAGTCACTATAATCTTTGGATCCAAACAAAGAAATTCGGAATTTCTGACGATAAT 1005
Db 85241 ATAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 85300
QY 1006 GTCGACTCAATTCATCTGAAAAAGGAGAGTTTGGTAGCTCATATCCCTTATATAT 1065
Db 85301 GATAATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 85360
QY 1066 GTACCTCTAAAGTAACTTAAAAACACACACCTTTCTTAAACCTATATAATAACCCCTA 1125
Db 85361 ACACAATAAATAATTTTATTTATATATTAATAATAATAATAATAATAATAATAATAATA 85420
QY 1126 A-TGGCTCACAAACGGGATAGTAGTACCCAGCTGTTTTTTTTTGGCTTATTTTT 1183
Db 85421 ATTTTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 85480
QY 1184 ATGACTACATTTTGTTCACITTTTGTTCGACITTAATACCGTTTATACCGTTCTCTT 1243

Db 85481 AITAAATATATATCTTTTATATATTTAAATTAATAATAATAATAATAATAATAATA 85540
QY 1244 TGATACACCTGATCGCGCTTTTAAACATAGCAACTCTTGTAAAGTCCCTTTCTTTTCC 1303
Db 85541 TTTATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 85600
QY 1304 CACTATTTTATCATCTTGAATAATATGTAATCAGATAGTTTTTCAAAAACATAATAAC 1363
Db 85601 TATTACTATATATTTAAATTAATAATAAGATATAAAATATCTTCTTAAATTTATAAT 85660
QY 1364 GGTCAAAATAACCGCTATTTTCAATTTTCCATTTCACTTTGTTTCTCAACAATATCAAC 1423
Db 85661 AATAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 85720
QY 1424 ACAACAGGAATCTCTATAGTCACCTCGCTTTTGTAGTT 1459
Db 85721 ATATTATATAATTTATATATTTTGTGTTAATAATTT 85756

RESULT 8

AC116921/c

LOCUS

DEFINITION

AC116921

AC116921

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AC116921 104014 bp DNA linear HTG 03-APR-2002
Dictyostelium discoideum chromosome 2 map 4603055-4707067 strain
AX4, *** SEQUENCING IN PROGRESS ***, in ordered pieces.

AC116921 GI:19909379
HTG: HTGS_PHASE2.
Dictyostelium discoideum.
Dictyostelium discoideum.
Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
1 (bases 1 to 104014)
Gloeckner, G., Eichinger, L., Szafranski, K., Pachebat, J., Dear, P.,
Lehmann, R., Baumgart, C., Parra, G., April, J., Guigo, R., Kumpf, K.,
Tunggal, B., Cox, E., Quail, M.A., Platzer, M., Rosenthal, A. and
Noegel, A.A.

Sequence and Analysis of Chromosome 2 of Dictyostelium
Unpublished
The Dictyostelium Genome Sequencing Consortium
2 (bases 1 to 104014)
Baumgart, C.
Direct Submission
Submitted (03-APR-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
CDS predictions from GenEd may contain errors. Further Information
is available from IMB Jena, Department of Genome Analysis
(http://genome.imb-jena.de/dictyostelium/)
and the University Cologne, Institute for Biochemistry I
(http://www.uni-koeln.de/dictyostelium/project.shtml)
Funding
Agency : Deutsche Forschungsgemeinschaft (DFG).
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

Location/Qualifiers
1. 104014
/organism="Dictyostelium discoideum"
/strain="AX4"
/db_xref="taxon:44689"
/chromosome="2"
/map="4603055-4707067"

BASE COUNT 40558 a 11425 c 11528 g 40251 t 252 others

ORIGIN

Query Match 5.1%; Score 75.2; DB 2; Length 104014;
Best Local Similarity 42.4%; Pred. No. 0.0029;
Matches 537; Conservative 0; Mismatches 718; Indels 11; Gaps 2;

QY 164 AATAACAGAAATACAGGAACCCCTCCAAAAAATTTTGGACCTTACACCCACATAA 223
Db 48756 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 48697

QY 224 ATTGGGATTAACCTTGGCCATAATAAAACCTCTTTGAACATACAGATATGTTATCTTTTC 283
D 48696 ATTTGAAAAAATAAAAAAGAGAGAGAAATAAAAAAATAATATATAGTCAACT 48637
QY 284 ATAACCTGGAATATTTTTCCTTTTAAACATATAGCAAAATTTGAAAAAATAAGGAAATG 343
D 48636 ATTCCTCTTTTGTGCTGTTGATTTGTAATAAATAATATGGGGAATAAAAAA 48577
QY 344 AAAAGGTAAAGTTCCTCAACCTGGAATATAGCTAAGGTTTTCCTGATCGCTT 403
D 48576 ATAAGTTAAAAAATTTTGAATAAATAAATAAATAAATAAATAAATAAATAA 48517
QY 404 AACTAAAGGAATAACAAAGTTATTAGGATAACCTGCGTAAGCTGCAACAAATA 463
D 48516 ATAAAAAGTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 48457
QY 464 TATTTGACCTTATAGCTCTATAGAAAAATATACAACTAAATCCCTTAAGAAATTTCCCTCA 523
D 48456 TCTTTAAAAATTTGAGTGGTGAATTTGAATCAAGATATAATAGGGGCTATTTATATA 48397
QY 524 TATATAATAGGAATCCCTCTCAGTGAAGTATCCATCTGAATATCAGTCCAC 583
D 48336 TTTTATTAATAAATTTTATTAATAATTTTATTAATAATTTTATTAATAAATAAATAA 48337
QY 584 TAATCCATCAATAAATAAGATTTAGTATTTGTTCTCTTCAGTACAATTTACTACATTTAT 643
D 48336 ATAAAACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 48277
QY 644 GCAATGCTAGCTTATTTGCTAATAATAGCAATTTGACACCCCTTAATTCGAACTTAAC 703
D 48276 TGAATTAAGTTGTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 48217
QY 704 GTATCCATATTTTCT 763
D 48216 TTTTATTTTCT 48157
QY 764 TAAAAAATTTTGAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 823
D 48156 AATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 8097
QY 824 TTTTGAAGTAAGAAATCAAAATTTTCTACACGACTATGCACTTACATCAACTGG 883
D 48096 GAATGTGTGAACGAACTCACTTTTATTAATTTTAAATTTTCGGTATTAATTAATTTA 48037
QY 884 ATGTATTTTGCATCTACTACTATAAGCTCA-----AACAAATTTCTTTCAAAATGTTAT 939
D 48036 TTATTTATTTTATTTATTTCTAAATAATTAATATATATATCTAGTAATAATTAATTTG 47977
QY 940 AATTACAAAGTCATCTATAATTTCTTTGGATCCAAAAACAAGGAATTCGGAATTCGAC 999
D 47976 TGATAATATTTATATAATAATACAATGTTATGGTGTCTTTGGTGGGAAATAATTTGA 47917
QY 1000 ATAAATGTCGACTCACAATTCATTTGTAAGAGGAGAGTTTGGTAGCTCATATCCCT 1059
D 47916 AAAAAAAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 47857
QY 1060 TATAATGTACCTCTAAAGTAATCTAAACAACAACACCTTTCTTAAACCTATAATA 1119
D 47856 AATCCCTTTGTCAGGAATTTTGTGAATAAATAAATAAATAAATAAATAAATAAATAA 47797
QY 1120 ACC-----TAATGGCTCAACACGGGATTAAGTTAGTTAGCCAGCTGTTTTTTTT 1172
D 47796 AACCGTTTTTGTGATGTCATATAAATTTGGAACCTGGAAGAGAAATTTCTATTTTT 47737
QY 1173 GCCTTATTTTATGACTACATTTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1232
D 47736 TTTTATTTTATAGATTTTTCATTTTTCATTTTTCATTTTTCATTTTTCATTTTTCAT 47677
QY 1233 CAACCTTCTCTTGTATCACTGCTATCCGCTTTTAAACATAGCACTCTCTGTAAGTCC 1292
D 47676 AGATTTTTTTTTTTTTTTTTTATATAGATAATTAATAAATAAATAAATAAATAA 47617
QY 1293 CTTTCTTTTCCCACTATTTTATCACTTCTGAAATATGTAATCAGATAGTTTTTCAAAA 1352

D 47616 CATTCATTAATAAATTTTTCATTTATTTTATTTTATTTTATTTTTCACAC 47557
QY 1353 CTATAAATAACGGTCAAAATACCGCTATTTCATTTTCCATTCACACTGTTTCTCAA 1412
D 47556 ATGTTAATAATAATAGCAACACCCCAATTTACGATTTTATTTTATTTT 47497
QY 1413 CAATAT 1418
D 47496 CAATTT 47491
RESULT 9
AE001398 14867 bp DNA linear INV 16-APR-2002
LOCUS Plasmodium falciparum chromosome 2, section 35 of 73 of the
DEFINITION complete sequence.
ACCESSION AE001398
VERSION AE001398.1 GI:3845197
KEYWORDS Plasmodium falciparum 3D7.
SOURCE Plasmodium falciparum 3D7.
ORGANISM Plasmodium falciparum 3D7.
AUTHORS Eukaryota: Alveolata; Apicomplexa; Haemosporida; Plasmodium.
Gardner, M.J., Tettelin, H., Carucci, D.J., Cummings, L.M., Aravind, L., Koonin, E.V., Shallow, S., Mason, T., Yu, K., Fujii, C., Pederson, J., Shen, K., Jing, J., Aston, C., Lal, Z., Schwartz, D.C., Pertea, M., Salzberg, S., Zhou, L., Sutton, G.G., Clayton, R., White, O., Smith, H.O., Fraser, C.M., Adams, M.D., Venter, J.C. and Hoffman, S.L.
Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum
Science 282 (5391), 1126-1132 (1998)
JOURNAL MEDLINE 99021743
PUBMED 9804551
REFERENCE 2 (bases 1 to 14867)
AUTHORS Gardner, M.J.
TITLE Direct Submission
JOURNAL Submitted (02-NOV-1998) The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20814, USA
FEATURES
Location/Qualifiers
1. .14867
/organism="Plasmodium falciparum 3D7"
/strain="3D7"
/db_xref="taxon:36329"
/chromosome="2"
complement(1570..2424)
/gene="PF0490c"
complement(1570..2424)
/gene="PF0490c"
/note="predicted by GlimmerM"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAC71887.1"
/db_xref="GI:3845198"
/translation="MKKEKIMDYLSCLDQVDREKKSGKNSLLKSSSTKSKDYK
SSIFSRRSHKSGSFRBSGFNRKSGSPKPYNNRLINKNNYKGNFNGSR
DNFKGTGSGFVDFNRKSGPKKREISNRKNSVSKYSGNGSKMRKSPKAPT
TVYTKLNNYKTVSAPVKNFNINISLYNRKNTFALNRKSPVGTIKSVPRKRIK
GLKSGSKSRKTRKSTGSKRFLNKIPKIVTSLNKIPSLPKEQKNTEVNLPSL
NNATTKN"
11241..14606
/gene="PF0495w"
11241..14606
/gene="PF0495w"
/note="predicted by GlimmerM"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAC71888.1"
/db_xref="GI:3845199"
/translation="MFALKNTVREGFVNICPSYKLYLKSNTVNLNLTNNK
TNKIKYKSKAQSLDFKGNINHLKLNPKYKCAKCEISAKVYKLLDEYK
FNTISLIDITQSKIFDEKTDYFNIEVKNDKNNKILFYFNKNDITFHR
EILGKCNKIMSYIHEMNGNELHFLIYFFRNKNDKNLILFYFNFYVDFHMYLFN

164	QY	164	ATATACAGAAATACAGGAACCCCTCAAAAAAAAATTTTGGACCTTTACACGCACATAA	223
7716	Db	7716	ATTAAATTAATTAATAAATTAATAATATATATATTTATATAAATTTAATATTAAA	7775
224	QY	224	ATTGGGATAAACCTTGCATTAATAAAAACCTTTGGAACATACGATATGTTATTCITTTTC	283
7776	Db	7776	ATATAATTATTAATTTATATAAATAATAAATAATAAATAATAATATATATATATTTTA	7835
284	QY	284	ATAACTGGAATPATTTTTCCTTTTTTAAACATTTAGCAAAATGGAATAAAAAAGGAATG	343
7836	Db	7836	ATTAATCTATTGTTATGTTAATAATAAACTATTATATTTAAATATATAAAATGTGA	7895
344	QY	344	AAAGGTGAAGAGTTCCTACCAATGGAATAATATAGGCTPAGGTTTTTCCTGATGCGCTTT	403
7896	Db	7896	TATAAATAAATAATCTTTTAAATTTATTTTTATAAATAAATTAATTTATTTAAATATT	7955
404	QY	404	AACTAAAAAGGAATAACAAAAGCTTATTAGCGATAACCTGCGTAAGGTGTCACACAAAATA	463
7956	Db	7956	ATTATATATATAAATTTATTTTATTTTAAATGTTAAATAATAAATAAATAAATAATA	8015
464	QY	464	TATTTTGCAGTTAGCTCTATAGAAAATATACAACATAAATCCTTTAAGGAATTTCCCTCTA	523
8016	Db	8016	AATTAATTAATAATAAATAAATAATAATTAATTAATTAATTAATTAATTAATTAATTA	8075
524	QY	524	TATATATATAGGAATCCCTCTCACGTGAACCTGAATTTATCCATCTGATTAATCAGTCCAC	583
8076	Db	8076	TATATATATATATTAATTTATTTATTTTAAATGAATAAATAAATAAATAAATAATAT	8135
584	QY	584	TAATCCATCAATAAAATAGATTAGTATGTTCTCTTCAGTACAATTAAGTACCATTAAT	643
8136	Db	8136	ATTAATTTAATTTATATAAATAATTTATTTATATATATTTAAATTAATTAATTAATTAAT	8195
644	QY	644	GCAATGCTAGCTTATGTTCCATAAATAGCCATGTTGCACACCCCTAAATTCGAACATTAAC	703
8196	Db	8196	TTTAAATAACTTAATTAATAATTAATTAATTTATATAAAAAAATAATATAAATTAATGATAT	8255
704	QY	704	GTATCCATATTTTCTTGCTCTTTGTTTGTCTTTTTC--TAACAAAAATGTTCCAGAAAT	760
8256	Db	8256	ATATATAATATAATATATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	8315
761	QY	761	TTTTAAAAAATVTTGAAAAACACATAACACTTTGGATGATGAATAATCAACTATTGAC	820
8316	Db	8316	TTTTTAATAAATATTATATATTTATTTATGAGTCTATTTTATTTTATTTTATTAATTAAT	8375
821	QY	821	TGTTTTGGAAGTAAAG--AATCAAAATTTTTTCTTAACCTGACCTAATGCACCTTTACATCA	878
8376	Db	8376	TAATTAATAATATATATTTAATAATAATATTTTGTTCTTTGATTAATAATAATTAATTT	8435
879	QY	879	ACTGGATGTTATTGGCATCTACTACTATAAGC--TCAACAAATATTCTTTCAAAAAATGT	936

REFERENCE 3 (bases 1 to 164399)
 LAWSON, D., Bowman, S. and Barrell, B.
 TITLE Submission
 JOURNAL Submitted (14-AUG-1997) P.falciparum Genome Sequencing Consortium,
 The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
 CB10 1SA, UK

COMMENT On or before May 14, 2001 this sequence version replaced
 gi:2665320, gi:2982570, gi:2982576, gi:2894503, gi:2982580,
 gi:2894373, gi:2982530.
 For more information about this sequence or the Malaria Project,
 see http://www.sanger.ac.uk/Projects/P_falciparum.

FEATURES
 source
 1..164399
 /organism="Plasmodium falciparum 3D7"
 /strain="3D7"
 /db_xref="taxon:36329"
 /chromosome="3"
 /clone="MAL3P6"
 complement(152..1075)
 /gene="PFC0700c, MAL3P6.1"
 complement(152..1075)
 /gene="PFC0700c, MAL3P6.1"
 /note="PFC0700c, MAL3P6.1", hypothetical protein, len: 307
 aa, predicted using hexExon"
 /codon_start=1
 /product="hypothetical protein, PFC0700c"
 /protein_id="CAB11150.1"
 /db_xref="GI:3758865"
 /db_xref="SPTREMBL:O77394"
 /translation="MHDFLLKSKFNLSPLFNFFYKRNDEYFKKRNNDLGLVM
 HNYADESEHREHKKDRMTSLKNEQIYTYNNFNNNYKYKSTKLEKRNED
 EYNEEOYEPTANLQDNKNDNMNFNNFNKNSLFNFQONADKNFLYLLKNNK
 NSTNENILVDPKLLKHWFLQMMNVNLOKOLLFNHLINTPKIMPHIIILNNKTEV
 SSNAVEINDDKKKNMVTLLKKILSSRENOMFVSSIIISFYLINKHQRALKI
 SQQKINSNFKLSVRLFEESLGIRKNYI"
 complement(156..161)
 /gene="PFC0700c, MAL3P6.1"
 /note="potential splice donor sequence at 3' end of
 PFC0700c may indicate splicing"
 2866..3762
 /gene="PFC0701w"
 2866..3762
 /gene="PFC0701w"
 /note="PFC0701w, hypothetical protein, len: 298 aa,
 revised: new gene prediction"
 /codon_start=1
 /product="hypothetical protein, PFC0701w"
 /protein_id="CAB90508.1"
 /db_xref="GI:7768284"
 /translation="MNTKLFNGSLCKILYKVINNLRGCYNIAQPFIIISYIKLEYI
 NNRKESCCSNIGIKKGYGNKMSPELLKRNIFNIYGNFCTMIKNMYNNINIY
 NSDYNVFKRYKWKVTSSYKRRKSSPNQOGOKAGIKRLSGEVVYTGOMLVK
 ORKICFENKTRRNRYFYGENYKVKAKNTSLIATNGRVKFFHVIQVNFVNLIL
 PEELELKQDLYRYTEHVYKFSPEENRSLIYURMKNLILFPYKIQFYIRPLKPOFL
 TKYDIYDNPFLQVPLLYHKRE"
 complement(join(4080..7204,7332..11478))
 /gene="PFC0705c, MAL3P6.2"
 complement(join(4080..7204,7332..11478))
 /gene="PFC0705c, MAL3P6.2"
 /note="PFC0705c, MAL3P6.2", hypothetical protein, len:
 2433 aa, predicted using hexExon but revised: changed
 splice donor sequence for exon 1"
 /codon_start=1
 /product="hypothetical protein, PFC0705c"
 /protein_id="CAB11149.2"
 /db_xref="GI:7768285"
 /db_xref="SPTREMBL:O77393"
 /translation="MTTFDIYKRLKKEKPCFKLRDSSPLSFQKCIIEWCDELIKYN
 IHDEKLLKLYRIKELIYSNLIYTKKNFOLPHMYNYKYNVYVROYMFLNSM
 KGDTYVSLKRYVGGKKKKIKYGVNGVYDNNNSHSDNNKNNKNDNNKDCDDN
 NNHNDNNNNHCDNNNNKHYVDNNNNKHYVDNNNNNNHCCVYFKRKEKNV
 DNIYSINIVDFDLVCKYKREKLYFMSKLYTCMIYCCITILKNIYIDFCLY
 LLKSYFFIDTIINKYIIVLPNPTCHKFLRLFLLLCNICNYGDTGENTYNEERIQ

DTNNICNDLYYYNYTNNIEDSYGKNSINNIYYHKKKIILNKKHFLIINKNDDNIN
 IGNIOFLSLFLYVPLKINFENMEFNFHTLQYSLCNLYLQFSQSYTRFLDLNKK
 VNNITFKCKEYFFKKNTLLTEKKNNNNIFLPMDYFPSPLOKRIKHNINFRHTNIH
 EHTTINNLVYHQSYSONMLNI PNSYENYENKLEKEDLNNLNTLFLSPLTIOQNN
 NNNKKYKYPHNLYGHNVLVDTKMKSLIFLLNHKETGPNFSLNREADQTSF
 NNNKKPHOFIFNNNYCYNKMSPHKNDIYDDDDFKNVLYIHYPTFKMNDIF
 TSSYNNITTKKKGNNQIILKNKNVNMKNIVNMKNVKNVKNVKNVKNVKNVKNV
 NKGKNNKTRTNTFQFLISPFKSENFVWKSFDVNTVNLKILENNIHTKKEKYN
 VFLTINSWYHAGNIINSNLNIRNFVFNRFNSNICNLAKYIYKNNVHTNDY
 NIWKIDILLCGLYGQNIEMSGRAHTEFRNNEKKGDKRWAKKKKKNKAVESIKG
 DEFOKDKNGLEKKGYPGSKSKKNNHMKKELIHSYCKKKKKKSECYISKD
 IICINSELAWLAKKNNKLCUYPKQNFQFCSMYQNYLHIFMYSKIGTFYRIKI
 FILVETIRNNRSTFKFKVRKSESDINLTAYEKKYKKNKMGNYKNIVRDKNI
 PLGEIFVYINCILLYIENKIRNIINDKVNPLHIYRIKKYGECEFLNCLCS
 CYINVKYRMFFKRIINFYKSIDNDISFSSFNNTCTNMKIRDYNNLSNIYGLSP
 YKNNNNNDMLYKNIINCLKKEFKKFLVNDONGVLYNKETIHELKPEOEFLRN
 KTCVNFSLFYPGNNELLYIYIYNLVNTOKKHLTKLCKELFLKMKPFLHFLT
 YVYMGINKDSFEYIINNEVAHEFFIFHNGKKYIDPKYLLSDVCISPLFNLSDV
 IYNTALLSSILRVAEEYFELSPKDSRENEKTOYRVDKTNMISITIRYNNYEN
 VDILNLFYTSMDKITSPQLYNSPOYKTKGKNDYENFIRYNDKECEVAYEKI
 KLDINKLEKDLILFYNNLNVHVMCLYDKRAWENYFEKKKIDPKKNVDKIYEY
 IRKKNKKNRINKYSYNNAKEYILKDLRCKOKIKKNNKILIKNNKNNYNNIY
 SRILNLFNKKKRNKLFSSVYVNLDCUHLHRTIRKHLIKNNFYNNKQNNKIQVDIK
 ELYSLSSVHNIIRDIECYQERKEMKDFDIIDHKEKLEINSENFVCDNISLGHNN
 NNNKALENGKSGAYRPGSNKNINKNINKNINKNINKNINKNINKNINKNINKN
 VNIQNNKNINININIDKGSNNPCNHNKREDIHYNNHYKHYFHNKHNDK
 SELLKNGGLEYNHEERKYEKISCKRNFENIENFYIRYIRYIRYIRYIRYIRYIR
 LINSFLNNMLLYYLLKCLPDDDDNYNFCNLILFNKRNKIINKIENDYNDNCC
 DHIFNKNHMQVFLNKNFYTSNKNKVDKNNLYTIDISTIFFIHEKIANPKLL
 RIFPDHVVYNDILYRLTSLYLSLNLNLIPLFSTLCRQYIFQCKKGGNPSTA
 EQGNIYQKNNMATLKNKLYKKEKNINFINKNSYKDDYKNDYKNDYKNDYKNDY
 ENLVNIOFKEDKPYMHIIININIDHLENNEDYINSCPYITKNTSNVYFMTEKYV
 NVGFKNFVHEKLFQNLKIYNDTILNDLYKIRSEMHLINALYDYMMLLSKRY
 PTFQKIVTHNEVDLKIHESEFVGHIFKLSFKTDLSPFIKIFININIVNIFKYIM
 KSIIPLGNNIKQLNVSKTCLNTYDEHTLLNVECKOHDYKSPFLFEQNEKLEHN
 IILTSORTQDFIKQFIYNNRHHFIQYVKNFNSDLSTIYQKFFENYIYHLCDESDYT
 P"

complement(7326..7331)
 /gene="PFC0705c, MAL3P6.2"
 /note="revised splice donor sequence cag/gtgcag for
 PFC0705c"
 join(14251..14913,15153..15215,15458..15874)
 /gene="PFC0710wMAL3P6.3"
 join(14251..14913,15153..15215,15458..15874)
 /gene="PFC0710wMAL3P6.3"
 /note="PFC0710w (MAL3P6.3), Inorganic pyrophosphatase,
 len: 380 aa: Similarity to inorganic pyrophosphatase, eg
 C.elegans (tr:Q18680) BLAST Score: 431, sum P(1) =
 9.2e-41; 43% identity in 201 aa overlap, predicted using
 hexExon, Pfam: match to PF00719 Pyrophosphatase, Inorganic
 pyrophosphatase Score 91.89"
 /codon_start=1
 /product="putative inorganic pyrophosphatase"
 /protein_id="CAB11148.1"
 /db_xref="GI:3758863"
 /translation="MGSKLINVEGNNQDDNNKSNVYISNNKVNNDYFIETNKEK
 KINLFONNIIISIN
 DNFPISPHHIDLKNDDGTVMIVETKYNIKLEQLREKFNVIKODKKGKGLRYV
 INSIYVYCALPQYIEPKHIYQNKSKKNKALLFTGNDPLDILDGACLKGQVY
 PVKILGATFLIDEGELEDWKIITAIKNDKHYEDINSLSDIEKYIYPTLSLLEWFRSYK
 MADYTKNLISQLYDKSEDLIMTKHHYLFREFEDVKYKLEHSEKTEKHEHYVNA
 QNTQFNVDKLNNDDEPMNNLLEDINITYYKSDSYAKPDLNIWTP"
 join(14620..14913,15153..15215,15458..15652)
 /gene="PFC0710wMAL3P6.3"
 /note="Match to PF00719 Pyrophosphatase, Inorganic
 pyrophosphatase Score 91.89"
 complement(16357..19383)
 /gene="PFC0715c, MAL3P6.4"
 complement(16357..19383)
 /gene="PFC0715c, MAL3P6.4"
 /note="PFC0715c (MAL3P6.4), hypothetical protein, len:
 1008 aa, predicted using hexExon"

misc_feature

gene

CDS

misc_feature

gene

CDS

```

/codon_start=1
/product="hypothetical protein, PFC0715c"
/protein_id="CAB1147.1"
/db_xref="GI:3758862"
/db_xref="SPUREBL:077391"
/translation="MTLYIFGIIVIIICSRFCNNVNNNIGKSHYPIFLKERNRNN
NFCSNYYNNGIRKKIKKNNKVPYSLGYIFLYNKLRSNTNNNNRNKCLVDYVP
KKKQGGNTIFLNHVIVINKMKQFNKVSICAEKSKRKRKRNNVHFMKDLGDI
FESNKTGDIIDKNSNVIINDSEYNHPIHNEGYNINIDENPNWYEGMPNHLSDSN
LESAMKADEYIKGKEDDIIEERKEKKKIVEMKDDDEFTDKLLDLIDYMLKYV
YRNIVYTSHELVLNIMCMKSPSYSNYVDNKEIQREKKKEEKKDEEKKKEE
LINSNEDQTHNETININIKRVNNDKOVNHEIDILGKKAKETDDEIQSEIEKD
ISNYSKDRNLFYTFSDKNNKLNEMSESLKYNKKNLKYFSLYKGNILKQLV
TEGEDLOPDEYTERKKDDTKYDSNNYDSNKNYNNYNNYNNKLNENNDKRL
KDTLSOLFDEEYEDYIDKERYMSNVTIEFNIYDYSNVLNKNKLNENNDKRL
IDSHLQIKCLLTKMKLEKAIIFYIKNNLHFSTSDNPINKDDRNIVDVIINEFWIK
MELYLGNIYGINEMKMGITPNMENEHKNMELKKNYSDNPONNMITIGDPSNKGV
DNYTNEMINKLSDDKKDALKNSAEVESSKNKYNNVLLNNSONVSIANLDEQI
KNVNNKVNKNVKNVKNVKNVKNVKNVKNVKNVKNVKNVKNVKNVKNVKNVKNV
KELEIKSDFLLKLENNEMKYDPNHVMKDLYPOLSEHHKNTDKYFENLKREMSNKC
SRGYTDNIKKKOSMGYDICEIGRTYIYIWNENFSLYFFLRPVYIKKKHILIDI
DNFFSLITINKHNIKIDIFNHPINSSDSIWSLTDNEDNHFMEENEMQFPYDIINI
Query Match          5.0%; Score 73.4; DB 3; Length 164399;
Best Local Similarity 42.9%; Pred. No. 0.0051;
Matches 470; Conservative 0; Mismatches 621; Indels 4; Gaps 2;

QY 268 ATATGTTTATCTTTTCATACACTGGAATATTTTTCGCTTTTTCATACATATGAAACAATTG 327
D 145198 ATATATATATATATATATATTTCTTTTTCATACATATGAAACAATTG 327

QY 328 AAAAAAAGGAATGAAAGGTAAGAGTTGCGCTACACATGTAAGAAATAATAGGCTAAGT 387
D 145258 ATAGCAAGAGAAATCGACAAGTCAACAAGTGAAGAATGATGATAGTTACGGAGGATCT 145317

QY 388 TTTTCCTGATCGCTTAACTAAAGAGAAATACAAAGTATTAACGATACCTCGGTA 447
D 145318 ATATCCAGTCTAATATGAAGTAATGATACACATATATGCTTTTATATATAACACGCA 145377

QY 448 AGGTGTCAACAAATATATTTTCAGCGTTAGCTCTATAGAAATATACAAACTAAATCTCT 507
D 145378 CATATATATATATATATATATATATATATATATATATATATATATATATATATTTTCT 145437

QY 508 TAAGGAATTTCTCTATATATATAGAAATCCCTCAGAGTGAAGTGAATATACCATC 567
D 145438 CAGCATATGATAGATAATACTATGCTGATGCTTCTTAATGGAAGGCTTCATTAAT 145497

QY 568 TGAATTATCAGTCCACTAATTCCTCAATAAATAGATAGTATGTTCTCTTCAGTA 627
D 145498 GAAGACTTTATCGATATATAGATATGAGATC--ATTAGCATATTTTCCCTGTGAATA 145555

QY 628 CAATTACTACCAATGCAATGCTAGCTTATTTTCATATAGGATCTTGCACACCCCT 687
D 145556 AAGTTTATTTTATATATATTTTATATTTTAAAGTATTTTAAAGATCTTATTTTACATAA 145615

QY 688 AATTCGAACATTAACGTATCATATATTTCTGTCCTCTCTGTTTCTTCTTCTTCTTCT 747
D 145616 AATTAATAAAGAGTATATACATATTTTATCATTTTATTTTATAGTTTATATAATTTT 145675

QY 748 ATGTTCCAGAAATTTTAAATAATATTTTCAAAAAACACATACACTTTGAGTATCATAT 807
D 145676 GTGATCTTTTATTTTATTTTATTTTTCGTGACTTCCAAAAATATATATATATATAT 145735

QY 808 ATCAACTATGACTGTTGTTTGAAGTAAAGATCAAAATTTTCTTAACCTGCAATGCT 867
D 145736 ATATATATATATTTCACTTGAAAAAATAAATAACATATATATATATATATATATATA 145795

QY 868 ACTTTACATCAACTGAGTGTATTTGCACTACTACTATAAGCTCAAAATAATATCTTT 927
D 145796 TATATATATATATATATATATTAACAAAGTTGCAATAATAACAAATCTATGCAATTA 145855

QY 928 CAAAAATGTTAATTAACAAAGTCACTATATTTCTTGGATCCAAACAAACAGGAATCG 987

```

```

Db 145856 ATAGCATATTAATAGATAATATAAATCATATAATATTATTTTGGGTAAAAAATATT 145915
QY 988 GAAATTCAGCATAAATTCGACTCACAATTCATTTGTAATAAGGAGAGATTTTGGTAGG 1047
Db 145916 AAAAAAATAATTAATTAAGGAAATTAATTAATAAAGAAATTAATTAATAAAGTTGG 145975
QY 1048 CTCATAATCGCTTATAATGCTACCTCTAAAGTAACTAAACAAACACACACCTTTCTAAA 1107
Db 145976 ACATAAAGGGGAATGATCAAAATTAATAAATAAAGAAATATATATATATATATATAT 146035
QY 1108 CCTATAATAATAACCTCTAAAGTAACTAAACAAACACACACCTTTCTTTT 1167
Db 146036 CATACATATTTGCTGCTGTTTAAACAAACACTTACATTTTAATAATTTTATTTTTC 146095
QY 1168 TTTTTCGCTTA--TTTATGACTACATTTTGTTCACACTTTTGTGTCGACTTTAATAACC 1225
Db 146096 ATTTTCTTAAGCTTTTCTTTTGTGTTTCTTCTCATATTTCTCTTTGTTGTCAGCAGAC 146155
QY 1226 GTTTTTCGAACTTCTCTTGTATACCTCTGATCCGCTTTTAAACATAGCAACTCTTGT 1285
Db 146156 ATTTTCTCACTTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 146215
QY 1286 AAAGTCCCTTTCTTTCCCACTATTTTATCATCTTGAATATGTAATGTAATGTAATGTTT 1345
Db 146216 TCTCTCAATCTCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 146275
QY 1346 TCAAAACATATAAT 1360
Db 146276 ACATTATCTTAACT 146290

RESULT 11
PFMAL3P4
LOCUS Plasmodium falciparum MAL3P4, complete sequence.
DEFINITION AL008970 ALJ39179
ACCESSION AL008970.3 GI:7672212
VERSION HTG.
KEYWORDS Plasmodium falciparum 3D7.
SOURCE Plasmodium falciparum 3D7.
ORGANISM Plasmodium falciparum 3D7.
REFERENCE 1 (bases 1 to 113880)
AUTHORS Bowman, S., Lawson, D., Basham, D., Brown, D., Chillingworth, T., Churcher, C.M., Craig, A., Davies, R.M., Devlin, K., Feltwell, T., Gentles, S., Gwilliam, R., Hamlin, N., Harris, D., Holroyd, S., Hornsby, T., Horrocks, P., Jagsal, K., Jassal, B., Kyes, S., McLean, J., Moule, S., Mungall, K., Murphy, L., Oliver, K., Quail, M.A., Rajadream, M.-A., Rutter, S., Skelton, J., Squares, R., Squares, S., Sulston, J.E., Whitehead, S., Woodward, J.R., Newbold, C. and Barrell, B.G.
TITLE The complete nucleotide sequence of chromosome 3 of Plasmodium falciparum
JOURNAL Nature 400 (6744), 532-538 (1999)
MEDLINE 99376085
PUBMED 1048855
REFERENCE 2 (bases 1 to 113880)
AUTHORS Hamlin, N., Bowman, S., Churcher, C., Lawson, D., Quail, M. and Barrell, B.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 113880)
AUTHORS Lawson, D., Bowman, S. and Barrell, B.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-1998) P.falciparum Genome Sequencing Consortium, The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK
COMMENT On Apr 30, 2000 this sequence version replaced gi:4493924. For more information about this sequence or the Malaria Project, see http://www.sanger.ac.uk/projects/P_falciparum.
FEATURES
source
1. .113880
/organism="Plasmodium falciparum 3D7"
/strain="3D7"
/db_xref="taxon:36329"

```



```

Db 190 ATTATATATATATTTTATTTTATATAAATTTTATTTTAAATAAATTAATTAATCAATTTATAC 131
QY 1012 TCACAAATTCATTTGTAAGGAGAGTTTGGTAGCTCATATCGCTTATATAGTACCT 1071
Db 130 ATAAATATATATATTAATTAATTTATTTTATATAATTTAAATTTTATTTTCAATTAATAA 71
QY 1072 CTAAAGTAATCTAAA 1087
Db 70 TTATATAAATATCACA 55

RESULT 13
AC005505/c
LOCUS
DEFINITION
AC005505 192929 bp DNA linear HTG 12-AUG-2000
Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN
PROGRESS ***, 1 ordered piece.
AC005505
VERSION
KEYWORDS
SOURCE
ORGANISM
Plasmodium falciparum
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE
1 (bases 1 to 192929)
Hyman, R.W., Fung, E.L., Qin, F., Rowley, D., Mao, J., Tamaki, T.,
Kurdi, O.B., Conway, A.B. and Davis, R.W.
Plasmodium falciparum 3D7 chromosome 12
Unpublished
TITLE
JOURNAL
REFERENCE
2 (bases 1 to 192929)
Hyman, R.W., Qin, F., Fung, E.L., Conway, A.B. and Davis, R.W.
Direct Submission
Submitted (21-AUG-1998) Stanford DNA Sequencing and Technology
Center, Stanford University, 855 California Avenue, Palo Alto, CA
94304, USA
COMMENT
On Aug 12, 2000 this sequence version replaced gi:8810446.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 192929: contig of 192929 bp in length.
FEATURES
Location/Qualifiers
1..192929
/organism="Plasmodium falciparum"
/db_xref="taxon:5833"
/chromosome="12"
/clone="PFAC181"
/clone="3D7"
BASE COUNT 81142 a 17397 c 22111 g 72279 t
ORIGIN

Query Match 4.88; Score 70.4; DB 2; Length 192929;
Best Local Similarity 43.78; Pred. No. 0.015;
Matches 458; Conservative 0; Mismatches 581; Indels 9; Gaps 3;

QY 232 TAAATTCGCCATATAAACTCTTTGAACATACGATATGTTATTTCTTTTCATACTGG 291
Db 43730 TATAAATTTCAATACATAAATTTCTTAATATATAAATGTTCTTTTAAACATCTTTT 43671
QY 292 AATATTTTGGCTTTTAAACATATGACAAATTTGAAAAAAGAAATGAAAGGTA 351
Db 43670 TTTTCTTTTAAATATCTTTTCCACAGTAATATACCATATGAAAAAATACCTACAA 43611
QY 352 AGAGTTGCTTAACCATTTGAAATATATAGGCTTAAGTTTTCCTGATCGGTTTAACTAAA 411
Db 43610 TTTTCTTTTAAATATCTTTTCCACAGTAATATACCATATGAAAAAATACCTACAA 43611
QY 412 AGGAATACAAAAGTTTATAGGCATACCTGCGCTAAGGTGTCACAAATATATTTTGC 471
Db 43550 AGTGCAAAAAAAGAAAGAAAAAAGAAAAAAGAAAAAATATACAAACATTTATATA 43491

```

```

QY 472 ACGTTAGCTCTATAGAAAAATATACAAACTAAATCCTTAAGGAATTTCCCTCTATATATAAT 531
Db 43490 TAAATAATCTTATATAAATTTTAAATATATATAAATAATATATATATATATATATATAT 43431
QY 532 AGGAAATCCCTCTCACAGTGAACCTGAATTTCCATCTGAATTTATCAGTCACCTCAATTTCCA 591
Db 43430 ATTATATAGATATATATAGATAATAATGTTATATATATATATATATATATATATAT 43371
QY 592 TCAATAAAAATAGATTAGTGTATGTTCTCTTTTCAGTACAAATTTACTACCATTTATGCAATGCT 651
Db 43370 CCAATAATAAAGCATTTAT -CATTTTAAAAAAGAAAAAATTTGAAATAGTAAATAATATAT 43312
QY 652 AGCTTATTTTCATTAATTTAGCCATGTTTCACACACCTTAATTCGAACATTAACCTGAT -CCA 710
Db 43311 AATATATATCTATTTATATACATATGATGTAAGTAAAAAGCATTTATTTATATATATATTA 43252
QY 711 TATTTTCTGTCCTCTCTTTGTTTCTTAAACAAAAATGTTCCAGAAATTTTAAAAAAA 770
Db 43251 TTAATAATATATAAATTTTCATCTCTTTGATCACAATTTTCTTTTCTTTTAAATTAAT 43192
QY 771 TATTTGAAAAACACATAACACTTTTGAGTATGATAATATCAACTATTGACATTTGTTTGA 830
Db 43191 TAAACAAATCTAAATAAATAAATTTAAAAAAGAAAAAATTTAAAAATATATACAA 43132
QY 831 AGTAAAGAAATCAAAATTTTCTTAACTCGACTAATGCACCTTTACATCAACTGCGATGTTAT 890
Db 43131 AAAAAATATTAACCATATATATATATATATATATATATATATATATATATATATATATA 43072
QY 891 TTGTCATCTACTACTATAAGCTCAAAACAAATTTATCTTTCAAAAATGTTTATAATTAACAACT 950
Db 43071 TTTTCTTTTATATAGAAACATGAATTTTCGAAGATAAATCTTTTCTTTTCTTCTTAAT 43012
QY 951 CATCTATAATTTCTTTGGATPCCAAAAACAAAGAAATTCGGAATTTCTGAGCATAAATGTCGA 1010
Db 43011 CTTTAAAAAATTTATATTTTATATATATATATATATATATATATATATATATATATAT 42952
QY 1011 CTCACAAATTCATGTTAAAAAGGGAGAGTTTGGTAGGCTCATATCGCTTATAATGATACC 1070
Db 42951 AACATTTTCTTTTATTTGAAAGCATATTTTATATATAAATATTA -TTTAAAAATAAA 42899
QY 1071 TCTAAGTAATCTAAAAACAAACACACACCTTTCTTAAACCTATATAATAACCTTAATGGC 1130
Db 42898 TAGAATAAAGAAAAAAGAAAAAAGAAAAAATATATATAATAATAATATATATATAT 42839
QY 1131 TCACAAACGGGATAAGTTAGTTAGCCAGCTGCTTTTTCCTGCTTTATTTTATGACTA 1190
Db 42838 ATAATATATATTTATATATACCATTAACACATTTATTTGTTATTTATACATATATATATA 42779
QY 1191 CATTTGTTTTCACCTTTTGTGCGACTTTAATACCGTTTTCGCACTTCTCTTTGCTATCA 1250
Db 42778 TTTATTTATTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTA 42719
QY 1251 CCTGTATCCGCTTTTAAAAACATAGCAA 1278
Db 42718 ATTTTATTTCTACATTTTCTGTAATAA 42691

```

```

RESULT 14
AC115598
LOCUS
DEFINITION
AC115598 56099 bp DNA linear HTG 21-MAR-2002
Dictyostelium discoideum chromosome 2 map 710700-766797 strain AX4.
*** SEQUENCING IN PROGRESS ***, in ordered pieces.
AC115598
VERSION
KEYWORDS
SOURCE
ORGANISM
Dictyostelium discoideum.
Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE
1 (bases 1 to 56099)
Lehmann, R., Baumgart, C., Parra, G., April, J.F., Guigo, R., Kumpf, K.,
Tunggal, B., Cox, E., Quail, M.A., Platzer, M., Rosenthal, A. and

```


Noegel, A.A.
 Sequence and Analysis of Chromosome 2 of Dictyostelium
 Unpublished
 The Dictyostelium Genome Sequencing Consortium
 2 (bases 1 to 56099)
 Baumgart, C.
 Direct Submission
 Submitted (21-MAR-2002) Genome Analysis, Institute of Molecular
 Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
 CDS predictions from GenBank may contain errors. Further Information
 is available from IMB Jena, Department of Genome Analysis
 (<http://genome.imb-jena.de/dictyostelium/>)
 and the University Cologne, Institute for Biochemistry I
 (<http://www.uni-koeln.de/dictyostelium/project.shtml>)
 Funding
 Agency : Deutsche Forschungsgemeinschaft (DFG).
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES
 source
 1..56099
 /organism="Dictyostelium discoideum"
 /strain="AX4"
 /db_xref="taxon:44689"
 /chromosome="2"
 /map="710700-766797"
 complement(join(62..712,782..1107,1156..9468,9524..14728,
 15111..15393))
 /note="ORF_ID:dd_03268"
 /pseudo
 /codon_start=1
 join(16701..16733,16844..16935,17099..17161,17502..17521,
 17648..17862)
 /note="ORF_ID:dd_00399"
 /pseudo
 /codon_start=1
 join(18663..19037,19129..19698)
 /note="ORF_ID:dd_00398"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="AAL92314.1"
 /db_xref="GI:19570018"
 /translation="MSINKEKTFNFRPLRTPKTKKSCSTHNEYSDLKIPKADGVPI
 ELNLTFFEYRIVKTRKIIISLMTDKLYITFAMGIAAFFLYFELFPRRLKI
 LAMIVFTLLVGLALSRITIMKLKAKLDOKIQILNDFKKNIIIBREYKEQVSS
 MPTKNHLLNIYLIQIADQNSKAESEKKEIKELQVIEKPIDKLIKNNYNS
 NDNNSSSTQLNISKLPQSQPIEIKIKNNNDYQKIEYSIEIIFKPNVNLN
 VDDNDNDVEIEMSMKANFEMVSNISYSDDKAPLIV"
 complement(19995..22391)
 /note="ORF_ID:dd_00396"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="AAL92315.1"
 /db_xref="GI:19570019"
 /translation="MNFYFPNKKSFCEIETVDNDEFNNNSRNSKATTIDSFINTPI
 IYDDYDEDDHEDIEGOEKKIKENKTIFCINPIYVINFKCGKIFSPNFHLLKI
 ATSKNDNWKIIEHYDLAVIPYKDGWVTVISDMYSTQROVDYFLKHLKQTN
 NITNGVELDGOQENIINKWTVENGKRFVGLFVIGDGDIEIRHNKPLFEK
 ISISGMEFDNKSISKYCLPNDITVINSNNRNPYFGLLEFDMFTDPFKMRL
 GSPSIESCFRANMKWVIEFDKLIYKGNQIGIILNDLYTSDWIEKLSGMDLYMD
 SLINHYTNKWSLIKSHNIIFNVYDINEINLGNVLLFNSONLRNHGIFRSCNLH
 RSEYSSLVFVIIITNNSNVSNPSFTQIPYIVSKOKLNKHVETKAVNSFKL
 SNLLTKLWNSYDYSKGEQFESQSLHLVCKHNSFLLSKIFSNLVGPKNYIQ
 SYLNFLNGLMCLONETPVYSYDMVQNTLYIKLFIYFPGLEFQNNPKSSNO
 NEYLTQVFFLQDQQRQTVYGYOTNATLYDNEYLYLNFNTSIEISNKKQVE
 IQNLKLLKLEIIEINDKSTVALTKFLYSNNELCGSDLVNMLNLFQNNFNEP
 QLNSTYASKKVKKEFDKIQIIDGIVNYTFLPQTIEPLITTSVTVTTCIR
 TVTVEITIKTTEIKSTVLKFNPTESCLTPILYVNVKRTQKSYENVCFSAISICRK
 "
 join(23293..23308,23497..24845)
 /note="ORF_ID:dd_00394"
 /codon_start=1

CDS
 complement(join(27712..28920,29055..29483))
 /note="ORF_ID:dd_00389"
 /codon_start=1
 /product="putative GTPase activating protein contains Pfam
 profile"
 /protein_id="AAL92318.1"
 /db_xref="GI:19570022"
 /translation="MPGISLEVPEKRGFLNLRSDATQYISPTTIGDNNWRCVLKE
 GCLFYENKDEIEVLSDKQKKEIKDNCFOLTATVFGNGVNGENGPIVTRY
 LSTPPFEFEYFAHIVSLRGVIFLRLSNSSNTSLSPQVNIKKSNQISLPLSTV
 SPGSPRETTITTTTTEKSHSGFVTHKVGSGNGSPSVNKLTPNKINSLSAPGPS
 STSDSSADDMNGFKKGLGVGQVTHKQKRWVYLNOQESISYKSQBPQTSK
 EPKSIQIIFCSAKVSTNDKYTFQVLTDTKTLHFSCVSGSMNLITCTQSAQIS
 MEALIKYTGCGGGGGSHSRPNSNTSSSVNSNRLSMRLKEDDVEDGVOK
 MDENKSLTKLQEDNKYCDGCPPLWAINLGVICINGSVGRNLGVHLSV
 RSVTMDIWRNMIOFRTDGNKVNOLYEYNIQPPQKLTPTDSTMEERKQIRSKYE
 HKLFSPSSSTTSSNHFENPSLASVSNPSSIV"
 complement(join(29656..29660,30433..31027))
 /note="ORF_ID:dd_00403"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="AAL92319.1"
 /db_xref="GI:19570023"
 /translation="MPPPTOPPLSVNGLGTICKQSQSELENKASIKNRNSTG
 TPSISLNGSGSGSGSGSSSFNPLSSELKSGNNIINLSTGSSSHLSPH
 SPQVHNSINSSEIPKPKRASIFGSKSKDKEKEKEKEKEKEKEKEKEKN
 NNSAASHIVSDDLQFAIKLVKVVVKNNAVESADGG"
 join(35631..37465,37551..37939,38058..42868)
 /note="ORF_ID:dd_03292"
 /codon_start=1
 /product="Putative Transferase"
 /protein_id="AAL92320.1"
 /db_xref="GI:19570024"
 /translation="MNDNSNSNDNTNNTLTATKNSINNTTTTTPNNVNCILKT
 SGSLKASSVPFPFGSASLPSILSNQYKLAASNVMNIMLSIISAAQQQQQ
 VPPQQTTTTATATATPTPTTSSSSMLNPLPKVKESIEFLFRNRPKN
 IREYVNSLSILPLDLSNTNKENIEFYNNIDENNYILKLLNTINDTKRL
 VYVYNNQHQPLQNTSPNTPTSTTTTSTKPSKGRKERGRNKKDKINEOL
 QQOQLQLOLQOQDHGSSLYLFPSSLSFIPKQOQOQOQSQOQLNIDQWVRK
 QQLDDNDNFNRVLEYNLKLIKYFKVMYQDSIKTHIFDLYIQNAQISEQR
 QLOEKESKTSKPIPOLISLVPTVLSIGDILKIKYATPSPNLTPTTTTMEEE
 ENTILINFEKTTDVSIIIESNLYLFECLISVSNEPTIVPNTTATATATTS
 RTTAAATTTPTLSSSSSPASQSSLSSSSSAVIAPTVTGYIFKKNELF
 RKILALVYTVVQHQITLSTVFTVLKGYLNEDQIQKLLNTLFLKSDSEOVKLY
 LTLKLSVSVFDLHNQKLNNDISLPYKLLIMGSTSPFENGQIFKIFDIWIMGI
 KQDQPSDSLTKYLLDFDSFNVISILFNQNKLTQOYLAAGSTMTPLSLHQIDD
 RRKFETILLSDELINFTWVEKSCFIANRLKSPGSAQTFEIKELLLQUNKOR
 DFKKILHFMESMEKLIIFSTVNGSVLVQSLANKETQFFRHNVRCEOWFSIRYNL
 LKASILSSSTPDIIRHASLRVQDIQANRFVDSNTAHFELEFILLNALQQLNETE
 SLOGLSNWSDINLNNNNNNNNNNNNSTFTFNNNGINKYTLKLPMMKQVILRS
 KQKPEESISLLSVPPMLESNSISFPVLEQIILKSYLDISNFTVEVQFLQNYQIQ
 QSNLSIIYKDTFKTGLGSFYRGEEMEAHAFKAKTESQNHILQREQLGLGLNQIWTD

CDS
 complement(join(25709..26010,26138..26168))
 /note="ORF_ID:dd_00392"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="AAL92317.1"
 /db_xref="GI:19570021"
 /translation="WIVPKLIGLNELNSNDYVPGQQTPTKVLSSRKSGLPTSPKPK
 SQKGPMTPTLFSKRESKDKPSSITMEDSQYLYILCGCKMCEKRFQHYMTIKHLK
 KKKKFFSP"
 complement(join(27712..28920,29055..29483))
 /note="ORF_ID:dd_00389"
 /codon_start=1
 /product="putative GTPase activating protein contains Pfam
 profile"
 /protein_id="AAL92318.1"
 /db_xref="GI:19570022"
 /translation="MPGISLEVPEKRGFLNLRSDATQYISPTTIGDNNWRCVLKE
 GCLFYENKDEIEVLSDKQKKEIKDNCFOLTATVFGNGVNGENGPIVTRY
 LSTPPFEFEYFAHIVSLRGVIFLRLSNSSNTSLSPQVNIKKSNQISLPLSTV
 SPGSPRETTITTTTTEKSHSGFVTHKVGSGNGSPSVNKLTPNKINSLSAPGPS
 STSDSSADDMNGFKKGLGVGQVTHKQKRWVYLNOQESISYKSQBPQTSK
 EPKSIQIIFCSAKVSTNDKYTFQVLTDTKTLHFSCVSGSMNLITCTQSAQIS
 MEALIKYTGCGGGGGSHSRPNSNTSSSVNSNRLSMRLKEDDVEDGVOK
 MDENKSLTKLQEDNKYCDGCPPLWAINLGVICINGSVGRNLGVHLSV
 RSVTMDIWRNMIOFRTDGNKVNOLYEYNIQPPQKLTPTDSTMEERKQIRSKYE
 HKLFSPSSSTTSSNHFENPSLASVSNPSSIV"
 complement(join(29656..29660,30433..31027))
 /note="ORF_ID:dd_00403"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="AAL92319.1"
 /db_xref="GI:19570023"
 /translation="MPPPTOPPLSVNGLGTICKQSQSELENKASIKNRNSTG
 TPSISLNGSGSGSGSGSSSFNPLSSELKSGNNIINLSTGSSSHLSPH
 SPQVHNSINSSEIPKPKRASIFGSKSKDKEKEKEKEKEKEKEKEKEKN
 NNSAASHIVSDDLQFAIKLVKVVVKNNAVESADGG"
 join(35631..37465,37551..37939,38058..42868)
 /note="ORF_ID:dd_03292"
 /codon_start=1
 /product="Putative Transferase"
 /protein_id="AAL92320.1"
 /db_xref="GI:19570024"
 /translation="MNDNSNSNDNTNNTLTATKNSINNTTTTTPNNVNCILKT
 SGSLKASSVPFPFGSASLPSILSNQYKLAASNVMNIMLSIISAAQQQQQ
 VPPQQTTTTATATATPTPTTSSSSMLNPLPKVKESIEFLFRNRPKN
 IREYVNSLSILPLDLSNTNKENIEFYNNIDENNYILKLLNTINDTKRL
 VYVYNNQHQPLQNTSPNTPTSTTTTSTKPSKGRKERGRNKKDKINEOL
 QQOQLQLOLQOQDHGSSLYLFPSSLSFIPKQOQOQOQSQOQLNIDQWVRK
 QQLDDNDNFNRVLEYNLKLIKYFKVMYQDSIKTHIFDLYIQNAQISEQR
 QLOEKESKTSKPIPOLISLVPTVLSIGDILKIKYATPSPNLTPTTTTMEEE
 ENTILINFEKTTDVSIIIESNLYLFECLISVSNEPTIVPNTTATATATTS
 RTTAAATTTPTLSSSSSPASQSSLSSSSSAVIAPTVTGYIFKKNELF
 RKILALVYTVVQHQITLSTVFTVLKGYLNEDQIQKLLNTLFLKSDSEOVKLY
 LTLKLSVSVFDLHNQKLNNDISLPYKLLIMGSTSPFENGQIFKIFDIWIMGI
 KQDQPSDSLTKYLLDFDSFNVISILFNQNKLTQOYLAAGSTMTPLSLHQIDD
 RRKFETILLSDELINFTWVEKSCFIANRLKSPGSAQTFEIKELLLQUNKOR
 DFKKILHFMESMEKLIIFSTVNGSVLVQSLANKETQFFRHNVRCEOWFSIRYNL
 LKASILSSSTPDIIRHASLRVQDIQANRFVDSNTAHFELEFILLNALQQLNETE
 SLOGLSNWSDINLNNNNNNNNNNNNSTFTFNNNGINKYTLKLPMMKQVILRS
 KQKPEESISLLSVPPMLESNSISFPVLEQIILKSYLDISNFTVEVQFLQNYQIQ
 QSNLSIIYKDTFKTGLGSFYRGEEMEAHAFKAKTESQNHILQREQLGLGLNQIWTD

stevor; T-complex protein 1 epsilon subunit; telomere; var.
 Plasmodium falciparum 3D7.
 Plasmodium falciparum 3D7.
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 1 (bases 1 to 253305)
 Bowman, S., Lawson, D., Basham, D., Brown, D., Chillingworth, T., Churcher, C.M., Craig, A., Davies, R.M., Devlin, K., Feltwell, T., Gentles, S., Gwilliam, R., Hamlin, N., Harris, D., Holroyd, S., Hornsby, T., Horrocks, P., Jagels, K., Jassal, B., Kves, S., McLean, J., Moule, S., Mungall, K., Murphy, L., Oliver, K., Quail, M.A., Rajandream, M.-A., Rutter, S., Skelton, J., Squares, R., Squares, S., Sulston, J.E., Whitehead, S., Woodward, J.R., Newbold, C. and Barrell, B.G.
 The complete nucleotide sequence of chromosome 3 of Plasmodium falciparum
 Nature 400 (6744), 532-538 (1999)
 93376085
 10448855
 2 (bases 1 to 253305)
 Oliver, K., Bowman, S., Harris, D., Lawson, D., Quail, M. and Barrell, B.
 Unpublished
 3 (bases 1 to 253305)
 Lawson, D., Bowman, S. and Barrell, B.
 Direct Submission
 Submitted (17-DEC-1998) P.falciparum Genome Sequencing Consortium, The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK
 On or before May 14, 2001 this sequence version replaced
 gi:2982540, gi:2982541, gi:2982544, gi:2894453, gi:2894454, gi:2982551, gi:2982556, gi:2982558, gi:2982563, gi:2982564, gi:2982566, gi:2894491, gi:2982569, gi:2695974, gi:2894502, gi:2673766, gi:2894496, gi:2982577, gi:2894588, gi:2894589, gi:2982538, gi:2982539, gi:2894380, gi:4725992.
 For more information about this sequence or the Malaria Project, see http://www.sanger.ac.uk/Projects/P_falciparum.
 Location/Qualifiers
 1. 253305
 /organism="Plasmodium falciparum 3D7"
 /strain="3D7"
 /db_xref="taxon:36329"
 /chromosome="3"
 /clone="MAL3P7"
 458. 4060
 /gene="PFC0860w, MAL3P7.1"
 458. 4060
 /gene="PFC0860w, MAL3P7.1"
 /note="PFC0860w (MAL3P7.1), Kinesin-related protein, len: 1200 aa; similarity: to kinesin-related protein.
 D.melanogaster kinesin-like protein 67A (TR:P91945) BLAST Score: 664, sum P(2) = 1.9e-67; 28% identity in 707 aa overlap, predicted using hexExon, Pfam: match to PF00225 kinesin, Kinesin motor domain Score 311.54"
 /codon_start=1
 /product="kinesin-related protein"
 /protein_id="CAB39023.1"
 /db_xref="GI:4493964"
 /db_xref="SPTREMBL:O97277"
 /translation="MRKTKNLNPPFLANNYYKFNPPNIIICDEKERTKIYSTNH
 HPTNLTNYDHTLKICNTENNKAHLHSNKSQKLLNPSYQKKGTHRTIKN
 VYPNIGPESVENTINGEFTINGQYDALKNMSLNNYDHHNNNNNNKNSFLV
 NNNNNNNQDVLENNEPTISKNIYDSIYIPQINIKNIINSEINNNNNNNNNH
 NYTNHNNNNNSNPTSSVTSKNNNNLNTNAYSNKRVAVRIKPIGESEENIVS
 IFNKVLTIEKENEKCYLLSKQKQYVDFSDVNAVTOBEVFFOTAKPLIPHFV
 GLNCTVFAVGATGSKTFTMLDDKNQNGIIVLSLELFTIIEKKCRNIKVLMFSLFV
 YNETRDLLGKKEKTKLEQDQVAVKYSNICEIEVNNYEQAMLLINCYKRRMSPT
 RANKVSSRSRHALQIYYNEILDDNMNTISAKLCFVDLAGSFASATSNKGRFKE
 GSYIQSLALANLNSIASNRNISKVRKYRDSKTLHLKNSLEGKCLVVMYANIP
 STSQENSTNTKYAFRANIKLATVQNDNKSDIEKILKNENLQKQYDITLIGY
 TNLKEFFTIINOLYKQISCYKLIENISDNSSMELKODITWYDOLVKKMSDEVR
 KKVDSLQDLYQKQFLNLFDTLPLEKLNLYVINSKVDNDNKKSLLEMIFFKHNK
 VDNVFLVNEKVVDKNVNGVNDENVVKNVNGVNDENVVNDENVVNDENVVNDENVV
 VDNVFLVNEKVVDKNVNGVNDENVVKNVNGVNDENVVNDENVVNDENVVNDENVV
 AEGKSESHNNKNDIEDNDKDTIKDIIHNNNSDNDNDDEYQSANSPVESDIAVKEKK

misc_feature
 /gene="PFC0860w, MAL3P7.1"
 /note="potential splice donor at 3' end of PFC0860w may indicate splicing"
 4057. 4062
 /note="potential splice donor at 3' end of PFC0860w may indicate splicing"
 join(5457. 5577, 5689. 5787, 5923. 6077, 6189. 6743)
 /gene="PFC0865w, MAL3P7.2"
 join(5457. 5577, 5689. 5787, 5923. 6077, 6189. 6743)
 /gene="PFC0865w, MAL3P7.2"
 /note="PFC0865w (MAL3P7.2), RNA-binding protein, len: 309 aa; Similarity eg to C.elegans RNA-binding protein (TR:Q18318) BLAST Score: 378, sum P(1) = 3.8e-35; 37% identity in 297 aa overlap, predicted using hexExon, Pfam: match to PF00076 rrm, RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain) Score 58.01"
 /codon_start=1
 /product="putative RNA-binding protein"
 /protein_id="CAB39067.1"
 /db_xref="GI:4494008"
 /db_xref="SPTREMBL:O97318"
 /translation="MGLFKIRNIEKLEAEKLNIGNNSDHWDDQYRDSSYIYIGNLD
 NMLTEGDIVFSGYQPIDVNLVNETKGTSGYCFLSVADORSTILAVDNFNGYKL
 LERPLVDHILNRLPKYLKADAKNYKPTGAGQGIYVNVVESIKLSKVPDKTK
 NKSNEKKKLLDDELWALNFESIKDKIISPLIGHDEKSRHNEGKKEEDEDDED
 DDDDDDDSDVHKYKHKRKRSLTTTKKKKKKKKSHDRKDKHRRHNSHREK
 EKDKSHKRRHKSDIKYSKRSKRSSTSSSTDR"
 join(7801. 7928, 8084. 8234, 8392. 8583)
 /gene="PFC0870w, MAL3P7.3"
 join(7801. 7928, 8084. 8234, 8392. 8583)
 /gene="PFC0870w, MAL3P7.3"
 /note="PFC0870w (MAL3P7.3), putative elongation factor, len: 181 aa; Similarity to P. faicparum elongation factor 1-beta, (AAF27524) (276 aa), fasta scores: opt: 366, E(): 1.1e-15, 53.2% identity in 141 aa overlap, revised: shortened exon 2, Pfam: match to PF00736 EF1BD, EF-1 guanine nucleotide exchange domain Score 61.10"
 /codon_start=1
 /product="putative elongation factor"
 /protein_id="CAB39068.2"
 /db_xref="GI:8052274"
 /db_xref="SPTREMBL:O97319"
 /translation="MANTYDELXVPLSYIYILQNEGNTSKIDQANTKKPKKKEVINKS
 SLIIDIKPYGENTDLDELKLVKNITMEGLTNGKAKHKKTPFAGLFLQVSCVIVDDL
 VNTDELLETIENTGLDNEQLQKKQMDDEENDEDEDEIGLVQSAEIIISFNKL"
 8235. 8240
 /gene="PFC0870w, MAL3P7.3"
 /note="revised splice donor sequence for exon 2 of PFC0870w"
 8580. 8585
 /note="potential splice donor sequence may indicate further splicing of PFC0870w"
 join(11431. 20416, 20578. 20690)
 /gene="PFC0875w, MAL3P7.4"
 join(11431. 20416, 20578. 20690)
 /gene="PFC0875w, MAL3P7.4"
 /note="PFC0875w (MAL3P7.4), hypothetical protein, len: 3032 aa, revised: added 3' exon, possibly spliced at 5' end"
 /codon_start=1
 /product="hypothetical protein, PFC0875w"
 /protein_id="CAB39024.2"
 /db_xref="GI:8052275"
 /db_xref="SPTREMBL:O97278"
 /translation="MKDFLSAENNDNLINFLKYNSEBCLKNLYAIELKSEDFKNVN
 SLENDILNHNKSLNLYNLHTNINIKSSEFNIDLYKDYTENEFENKLNKI
 NPHLLKVVKNTEFYKNFIKDKNLKNENHYQYFNDNRKKLFYFNKKNLVETKYS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

PUBMED

AUTHORS

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

gene

CDS

